



Fourier-based solving approach for the transport-of-intensity equation with reduced restrictions

SOHEIL MEHRABKHANI,^{1,2,*} LENNART WEFELNBERG,¹ AND THOMAS SCHNEIDER^{1,2}

¹Terahertz-Photonics Group, Institut für Hochfrequenztechnik, TU Braunschweig, Schleinitzstraße 22, 38106 Braunschweig, Germany

²Lena, Laboratory for Emerging Nanometrology, Braunschweig, Germany

*soheil.mehrabkhani@ihf.tu-bs.de

Abstract: The transport-of-intensity equation (TIE) has been proven as a standard approach for phase retrieval. Some high efficiency solving methods for the TIE, extensively used in many works, is based on a Fourier transform (FT). However, several assumptions have to be made to solve the TIE by these methods. A common assumption is that there are no zero values for the intensity distribution allowed. The two most widespread Fourier-based approaches have further restrictions. One of these requires the uniformity of the intensity distribution and the other assumes the parallelism of the intensity and phase gradients. In this paper, we present an approach, which does not need any of these assumptions and consequently extends the application domain of the TIE.

© 2018 Optical Society of America under the terms of the [OSA Open Access Publishing Agreement](#)

OCIS codes: (100.5070) Phase retrieval; (050.1960) Diffraction theory; (110.1758) Computational imaging; (110.1650) Coherence imaging; (090.1995) Digital holography.

References and links

1. K. A. Nugent, "Coherent methods in the X-ray sciences," *Adv. Phys.* **59**(1), 1–99 (2010).
2. K. Yatabe, K. Ishikawa, and Y. Oikawa, "Simple, flexible, and accurate phase retrieval method for generalized phase-shifting interferometry," *J. Opt. Soc. Am. A* **34**(1), 87–96 (2017).
3. K. Creath, "V Phase-Measurement Interferometry Techniques," *Prog. Opt.* **26**, 349–393 (1988).
4. M. R. Teague, "Deterministic phase retrieval: a Green's function solution," *J. Opt. Soc. Am.* **73**(11), 1434–1441 (1983).
5. R. Amézquita-Orozco and Y. Mejía-Barbosa, "Gerchberg-Saxton algorithm applied to a translational-variant optical setup," *Opt. Express* **21**(16), 19128–19134 (2013).
6. C. Roddier and F. Roddier, "Wavefront reconstruction from defocused images and the testing of ground based optical telescopes," *J. Opt. Soc. Am. A* **10**(11), 2277–2287 (1993).
7. R. Yazdani, M. Hajimahmoodzadeh, and H. R. Fallah, "Application of the transport of intensity equation in determination of nonlinear refractive index," *Appl. Opt.* **53**(35), 8295–8301 (2014).
8. J. Martinez-Carranza, K. Falaggis, and T. Kozacki, "Fast and accurate phase-unwrapping algorithm based on the transport of intensity equation," *Appl. Opt.* **56**(25), 7079–7088 (2017).
9. W. J. Zhou, X. Guan, F. Liu, Y. Yu, H. Zhang, T. C. Poon, and P. P. Banerjee, "Phase retrieval based on transport of intensity and digital holography," *Appl. Opt.* **57**(1), A229–A234 (2018).
10. E. Bostan, E. Froustey, B. Rappaz, E. Shaffer, D. Sage, and M. Unser, "Phase retrieval by using transport-of-intensity equation and differential interference contrast microscopy," in *Proceedings of 2014 IEEE International Conference on Image Processing (ICIP) (IEEE, 2014)*, pp. 3939–3943.
11. P. Soltani, A. Darudi, G. Nehmetallah, A. R. Moradi, and J. Amiri, "Accurate testing of aspheric surfaces using the transport of intensity equation by properly selecting the defocusing distance," *Appl. Opt.* **55**(35), 10067–10072 (2016).
12. C. Dorner and J. D. Zuegel, "Optical testing using the transport-of-intensity equation," *Opt. Express* **15**(12), 7165–7175 (2007).
13. M. Basunia, P. P. Banerjee, U. Abeywickrema, T. C. Poon, and H. Zhang, "Recursive method for phase retrieval using transport of intensity and its applications," *Appl. Opt.* **55**(33), 9546–9554 (2016).
14. R. Shomali, A. Darudi, and S. Nasiri, "Application of irradiance transport equation in aspheric surface testing," *Optik (Stuttg.)* **123**(14), 1282–1286 (2012).
15. Z. Jingshan, R. A. Claus, J. Dauwels, L. Tian, and L. Waller, "Transport of Intensity phase imaging by intensity spectrum fitting of exponentially spaced defocus planes," *Opt. Express* **22**(9), 10661–10674 (2014).

16. N. Pandey, A. Ghosh, and K. Khare, "Two-dimensional phase unwrapping using the transport of intensity equation," *Appl. Opt.* **55**(9), 2418–2425 (2016).
17. S. C. Woods and A. H. Greenaway, "Wave-front sensing by use of a Green's function solution to the intensity transport equation," *J. Opt. Soc. Am. A* **20**(3), 508–512 (2003).
18. N. Pandey, A. Ghosh, and K. Khare, "Two-dimensional phase unwrapping using the transport of intensity equation," *Appl. Opt.* **55**(9), 2418–2425 (2016).
19. B. Xue, S. Zheng, L. Cui, X. Bai, and F. Zhou, "Transport of intensity phase imaging from multiple intensities measured in unequally-spaced planes," *Opt. Express* **19**(21), 20244–20250 (2011).
20. Z. Jingshan, R. A. Claus, J. Dauwels, L. Tian, and L. Waller, "Transport of Intensity phase imaging by intensity spectrum fitting of exponentially spaced defocus planes," *Opt. Express* **22**(9), 10661–10674 (2014).
21. T. Nguyen, G. Nehmetallah, D. Tran, A. Darudi, and P. Soltani, "Fully automated, high speed, tomographic phase object reconstruction using the transport of intensity equation in transmission and reflection configurations," *Appl. Opt.* **54**(35), 10443–10453 (2015).
22. L. Waller, L. Tian, and G. Barbastathis, "Transport of intensity phase-amplitude imaging with higher order intensity derivatives," *Opt. Express* **18**(12), 12552–12561 (2010).
23. D. Paganin and K. Nugent, "Noninterferometric phase imaging with partially coherent light," *Phys. Rev. Lett.* **80**(12), 2586–2589 (1998).
24. Y. Zhu, A. Shanker, L. Tian, L. Waller, and G. Barbastathis, "Low-noise phase imaging by hybrid uniform and structured illumination transport of intensity equation," *Opt. Express* **22**(22), 26696–26711 (2014).
25. C. Zuo, Q. Chen, L. Huang, and A. Asundi, "Phase discrepancy analysis and compensation for fast Fourier transform based solution of the transport of intensity equation," *Opt. Express* **22**(14), 17172–17186 (2014).
26. L. Huang, C. Zuo, M. Idir, W. Qu, and A. Asundi, "Phase retrieval with the transport-of-intensity equation in an arbitrarily shaped aperture by iterative discrete cosine transforms," *Opt. Lett.* **40**(9), 1976–1979 (2015).
27. C. Zuo, Q. Chen, and A. Asundi, "Phase retrieval and computational imaging with the transport of intensity equation: boundary conditions, fast solution, and discrepancy compensation," in *Classical Optics 2014*, OSA Technical Digest (online) (Optical Society of America, 2014), paper CT3C.3.
28. C. Zhang, W. He, J. Wu, and X. Peng, "Optical cryptosystem based on phase-truncated Fresnel diffraction and transport of intensity equation," *Opt. Express* **23**(7), 8845–8854 (2015).
29. T. Chakraborty and J. C. Petrucci, "Source diversity for transport of intensity phase imaging," *Opt. Express* **25**(8), 9122–9137 (2017).
30. A. Shanker, L. Tian, M. Sczyrba, B. Connolly, A. Neureuther, and L. Waller, "Transport of intensity phase imaging in the presence of curl effects induced by strongly absorbing photomasks," *Appl. Opt.* **53**(34), J1–J6 (2014).
31. M. Krenkel, M. Bartels, and T. Salditt, "Transport of intensity phase reconstruction to solve the twin image problem in holographic x-ray imaging," *Opt. Express* **21**(2), 2220–2235 (2013).
32. V. V. Volkov, Y. Zhu, and M. De Graef, "A new symmetrized solution for phase retrieval using the transport of intensity equation," *Micron* **33**(5), 411–416 (2002).
33. C. Zuo, J. Sun, J. Li, J. Zhang, A. Asundi, and Q. Chen, "High-resolution transport-of-intensity quantitative phase microscopy with annular illumination," *Sci. Rep.* **7**(1), 7654 (2017).
34. C. Zuo, Q. Chen, and A. Asundi, "Boundary-artifact-free phase retrieval with the transport of intensity equation: fast solution with use of discrete cosine transform," *Opt. Express* **22**(8), 9220–9244 (2014).
35. C. Zuo, Q. Chen, W. Qu, and A. Asundi, "High-speed transport-of-intensity phase microscopy with an electrically tunable lens," *Opt. Express* **21**(20), 24060–24075 (2013).
36. J. Martinez-Carranza, K. Falaggis, T. Kozacki, and M. Kujawinska, "Effect of imposed boundary conditions on the accuracy of transport of intensity equation based solvers," *Proc. SPIE* **8789**, 87890N (2013).
37. J. A. Schmalz, T. E. Gureyev, D. M. Paganin, and K. M. Pavlov, "Phase retrieval using radiation and matterwave fields: Validity of Teague's method for solution of the transport-of-intensity equation," *Phys. Rev. A* **84**(2), 023808 (2011).
38. J. Martinez-Carranza, K. Falaggis, and T. Kozacki, "Optimum phase retrieval using the transport of intensity equation," *Proc. SPIE* **9132**, 91320T (2014).
39. T. C. Poon and J. P. Liu, *Introduction to modern digital holography with MATLAB* (Cambridge University, 2014).
40. T. M. Apostol, *Mathematical Analysis*, 2th ed. (Addison-Wesley, 1981).
41. P. Picart and J. C. Li, *Digital Holography* (Wiley, 2012).
42. J. W. Goodman, *Introduction to Fourier Optics*, 4th ed. (W. H. Freeman, 2017).
43. K. Falaggis, T. Kozacki, and M. Kujawinska, "Optimum plane selection criteria for single-beam phase retrieval techniques based on the contrast transfer function," *Opt. Lett.* **39**(1), 30–33 (2014).
44. J. Martinez-Carranza, K. Falaggis, and T. Kozacki, "Multi-filter transport of intensity equation solver with equalized noise sensitivity," *Opt. Express* **23**(18), 23092–23107 (2015).
45. R. Bie, X.-H. Yuan, M. Zhao, and L. Zhang, "Method for estimating the axial intensity derivative in the TIE with higher order intensity derivatives and noise suppression," *Opt. Express* **20**(7), 8186–8191 (2012).
46. D. Paganin, A. Barty, P. J. McMahon, and K. A. Nugent, "Quantitative phase-amplitude microscopy. III. The effects of noise," *J. Microsc.* **214**(1), 51–61 (2004).
47. J. Martinez-Carranza, K. Falaggis, and T. Kozacki, "Optimum measurement criteria for the axial derivative intensity used in transport of intensity-equation-based solvers," *Opt. Lett.* **39**(2), 182–185 (2014).

48. M. Soto and E. Acosta, "Improved phase imaging from intensity measurements in multiple planes," *Appl. Opt.* **46**(33), 7978–7981 (2007).
49. J. Martinez-Carranza, K. Falaggis, M. Jozwick, and T. Kozacki, "Comparison of phase retrieval techniques based on the transport of intensity equation using equally and unequally spaced plane separation criteria," *Proc. SPIE* **9204**, 92040G (2014).
50. M. Soto, E. Acosta, and S. Ríos, "Performance analysis of curvature sensors: optimum positioning of the measurement planes," *Opt. Express* **11**(20), 2577–2588 (2003).

1. Introduction

Within the last decades, the number of applications for coherent diffraction has grown enormously. Among them are biomedical imaging, coherent tomography, digital holography, coherent X-ray imaging, holographic microscopy, optical metrology, wave sensing and many more. The information of the coherent illumination of an object is contained in a complex scalar field, composed of intensity and phase. Although, the intensity distribution can be measured by a detector, the phase measurement is a rather challenging task. As a standard technique, interferometry has been used [1–3]. However, it requires a cumbersome very sensitive experimental setup. Due to powerful computer processors, available in recent years, phase retrieval has been established as the most widely used alternative to interferometry. The main advantage of phase retrieval is its ability to extract the phase information from the intensity measurements by appropriate numerical algorithms. Until now, several phase retrieval algorithms have been developed. However, each method has its drawbacks and limitations, which restrict its application. One of the algorithms in the paraxial domain is based on the transport-of-intensity equation (TIE) originally developed by Teague [4]. Contrary to iterative approaches like the Gerchberg-Saxton algorithm [5] and its numerous modifications, the TIE is a deterministic approach. Thus, it avoids problems like non-convergence or stagnation, which are typical in many iterative approaches. Although, it is an elegant deterministic way for phase retrieval, its solution is not trivial. Two very successful fast solving approaches are based on the Fast Fourier Transform (FFT). However, they suffer from different assumptions related to the intensity behavior, which automatically restrict the application of these tools. One of these approaches proposed by [6] and applied in [6–24], for instance, assumes the uniformity of the intensity distribution (orthogonal to the propagation direction). Another approach [25–37], is also restricted to very special cases due to the assumption that the transverse gradient of the phase and the intensity must be parallel [37]. Moreover, zero values for the intensity must be strictly avoided in both approaches. In this work, we present an efficient algorithm without these restrictions, making the applicability of the TIE more general. We also present results under consideration of noise.

2. Iterative solution for the TIE

The TIE describes the coupling between the intensity I and the phase ϕ in a plane orthogonal to the propagation direction of the paraxial field [4]

$$\nabla_{\perp} \cdot (I \nabla_{\perp} \phi) = -k \frac{\partial I}{\partial z}, \quad (1)$$

where k is the wave number ($k = 2\pi/\lambda$). If the complex amplitude is replaced by, $\sqrt{I} \exp(i\phi)$ the TIE is the imaginary part of the paraxial Helmholtz equation [38]. In principle, the two coefficients k and I are known. The axial intensity derivative $\partial I / \partial z$ can be linearly approximated by the finite difference method using two intensity measurements in two parallel planes orthogonal to the propagation with the distance dz :

$$\frac{\partial I}{\partial z} \approx \frac{I_2 - I_1}{dz}. \quad (2)$$

In order to remove the aforementioned restrictions of the standard solutions of the TIE, we propose the following change of Eq. (1):

$$(I + C) \nabla_{\perp}^2 \phi = -k \frac{\partial I}{\partial z} - \nabla_{\perp} I \cdot \nabla_{\perp} \phi + C \nabla_{\perp}^2 \phi, \quad (3)$$

where the parameter C is an arbitrary positive constant, $C(x, y) = \text{const}$. Since it avoids the division by zero, the constant C enables zero values in the intensity distribution. It should be noted, that the Eqs. (1) and (3) are basically equivalent. Equation (3) is simply derived by adding $C \nabla_{\perp}^2 \phi$ on both sides of Eq. (1) and expanding the divergence operator. In this general form, Eq. (3) cannot be solved easily, however, an iterative approach may be applied to convert it to a Poisson's equation. Such an iterative Poisson's equation additionally avoids the restrictions of the other TIE solving approaches, mentioned above. Let us consider the phase on the right side of Eq. (3) as the phase after the n -th iteration, whereas the left side is the next one. This way, the phase on the left and right side can be treated as two different mathematical variables ϕ_n and ϕ_{n+1} in an iterative approach. Consequently, Eq. (3) can be written as a Poisson's equation:

$$\nabla_{\perp}^2 \phi_{n+1} = \frac{-k \frac{\partial I}{\partial z} - \nabla_{\perp} I \cdot \nabla_{\perp} \phi_n + C \nabla_{\perp}^2 \phi_n}{I + C}. \quad (4)$$

To solve Eq. (4), the Fourier Transform (FT) can be used:

$$-4\pi^2 (f_x^2 + f_y^2) \tilde{\phi}_{n+1} = FT \left\{ \frac{-k \frac{\partial I}{\partial z} - \nabla_{\perp} I \cdot \nabla_{\perp} \phi_n + C \nabla_{\perp}^2 \phi_n}{I + C} \right\}. \quad (5)$$

For all frequency pairs $(f_x, f_y) \neq 0$ Eq. (5) can be rewritten as:

$$\tilde{\phi}_{n+1} = \frac{1}{-4\pi^2 (f_x^2 + f_y^2)} FT \left\{ \frac{-k \frac{\partial I}{\partial z} - \nabla_{\perp} I \cdot \nabla_{\perp} \phi_n + C \nabla_{\perp}^2 \phi_n}{I + C} \right\}. \quad (6)$$

The singularity at the frequency point $(f_x, f_y) = 0$ may be simply removed by a constant predefined value for the Fourier transform of the phase at the frequency center $\tilde{\phi}_{n+1}(0, 0) = 0$. This corresponds to a constant phase shift in the spatial domain, which has no influence on the result. Equation (6) provides $\tilde{\phi}_{n+1}(f_x, f_y)$, which is the FT of $\phi_{n+1}(x, y)$. Finally, the phase $\phi_{n+1}(x, y)$ is calculated by the inverse Fourier transform (IFT):

$$\phi_{n+1} = \frac{-1}{4\pi^2} IFT \left\{ \frac{1}{f_x^2 + f_y^2} FT \left\{ \frac{-k \frac{\partial I}{\partial z} - \nabla_{\perp} I \cdot \nabla_{\perp} \phi_n + C \nabla_{\perp}^2 \phi_n}{I + C} \right\} \right\}. \quad (7)$$

The initial phase is set to $\phi_0(x, y) = 0$. To get the next estimate for the phase ϕ_{n+1} , in each iteration the phase ϕ_n is introduced into Eq. (7). The iteration will be repeated, until the

convergence is reached. The termination condition for the algorithm may be determined by a fixed iteration number, for instance.

3. Numerical Implementation

3.1 Constraints on the phase changes under the paraxial assumption

The TIE is only valid in the paraxial domain and consequently its solution must satisfy the paraxial condition. However, in the proposed iterative method, the starting phase distribution $\phi_0(x, y) = 0$ and subsequently calculated phase distributions $\phi_n(x, y)$ are not necessarily a solution of the TIE. They are just approximations for the correct solution, which must be improved in each iteration. Thus, these estimates may violate the paraxial assumption. Consequently, some a priori information concerning the paraxial domain must be included as additional constraints to prevent a non-physical solution.

According to the Fresnel transform (FRT), there is the following relationship between the position and frequency [39]:

$$\vec{f}_\perp = \frac{\vec{r}_\perp}{\lambda z} = f \frac{\vec{r}_\perp}{z}, \quad (8)$$

where $\vec{f}_\perp = (f_x, f_y)$ is the transverse frequency vector corresponding to the object (image) plane and $\vec{r}_\perp = (x, y)$ the transverse position vector in the image (object) plane. The λ and f are the wavelength and the spatial frequency of the wave respectively and z is the propagation distance between the object and the first image plane, Fig. 1. According to Eq. (8), the maximum transverse frequency corresponding to the image (with intensity distribution I_1) is $f_{\perp \max} = r_{\perp \max} / \lambda z \approx \sin(\theta) / \lambda$. Thus, with the definition of the numerical aperture $NA_{obj} = \sin(\theta)$, the maximum transverse frequency is $f_{\perp \max} = NA_{obj} f$. Consequently, the transverse wave vector component $k_\perp = 2\pi f_\perp$ satisfies the inequality:

$$k_\perp \leq NA_{obj} k. \quad (9)$$

It should be mentioned, that inequality (9) has been derived using Eq. (8), which is a consequence of the paraxial assumption used for the Fresnel transform. Consequently, inequality (9) is a result of the paraxial assumption as well.

Figure 1 shows the angle θ , which is the angle between the paraxial propagation axis (orthogonal to the image planes) and the line going through the center of the first image and the point of the object with the farthest distance to the paraxial axis. This angle determines the numerical aperture NA_{obj} and thus the maximum possible transverse wave vector component corresponding to the first image.

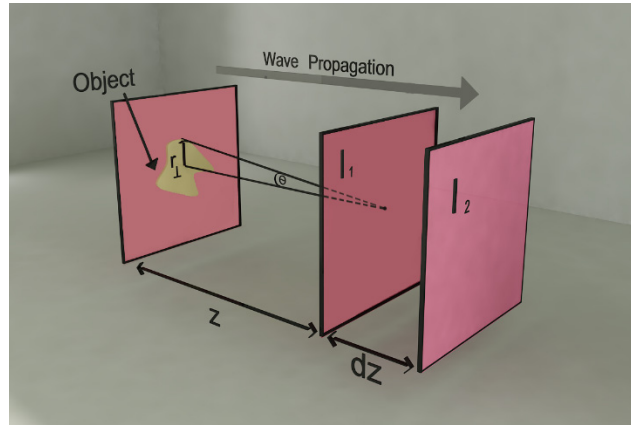


Fig. 1. Theoretical setup for the TIE with the object plane and the two intensity measurements in close proximity dz .

According to the definition of the local spatial frequency [39], the following relationships can be derived for the corresponding local wave vector components:

$$k_{1,x} = \frac{\partial \phi(x)}{\partial x}, \quad k_{1,y} = \frac{\partial \phi(y)}{\partial y}, \quad k_{2,x} = \frac{\partial \phi(x+h)}{\partial x}, \quad k_{2,y} = \frac{\partial \phi(y+h)}{\partial y}, \quad (10)$$

which, together with the inequality (9), result in the following paraxial condition for the transverse gradient of the phase:

$$|\nabla_{\perp} \phi| \leq NA_{obj} k. \quad (11)$$

The right side of Eq. (7) depends on the vector $\nabla_{\perp} \phi$ and the scalar $\nabla_{\perp}^2 \phi$ thus, the paraxial condition concerning $\nabla_{\perp}^2 \phi$ must be derived as well. According to the definition of the Laplace operator including its discretization with the spacing $\Delta x = \Delta y = h$, it follows:

$$|\nabla_{\perp}^2 \phi| = \left| \frac{\partial^2 \phi}{\partial x^2} + \frac{\partial^2 \phi}{\partial y^2} \right| \approx \frac{1}{h} \left| \frac{\partial \phi(x+h)}{\partial x} - \frac{\partial \phi(x)}{\partial x} + \frac{\partial \phi(y+h)}{\partial y} - \frac{\partial \phi(y)}{\partial y} \right|. \quad (12)$$

The triangle inequality [40] within the Eqs. (10) and (12) provide a very suitable estimate for the maximum of $|\nabla_{\perp}^2 \phi|$:

$$|\nabla_{\perp}^2 \phi| \approx \frac{1}{h} |k_{2,x} - k_{1,x} + k_{2,y} - k_{1,y}| \leq \frac{1}{h} (|k_{1,x} + k_{1,y}| + |k_{2,x} + k_{2,y}|). \quad (13)$$

Equation (10) and the inequality (11) determine a maximum for $|k_x + k_y|$:

$$|\nabla_{\perp} \phi|^2 \approx k_x^2 + k_y^2 \leq NA_{obj}^2 k^2 \Rightarrow |k_x + k_y| \leq \sqrt{NA_{obj}^2 k^2 + 2k_x k_y}, \quad (14)$$

which will be applied in inequality (13):

$$|\nabla_{\perp}^2 \phi| \leq \frac{1}{h} \left(\sqrt{NA_{obj}^2 k^2 + 2k_{1,x} k_{1,y}} + \sqrt{NA_{obj}^2 k^2 + 2k_{2,x} k_{2,y}} \right). \quad (15)$$

The advantage of inequality (15) is its dependence on the terms $k_{1,x} k_{1,y}$ and $k_{2,x} k_{2,y}$, whose maximum can be simply found. After, defining the function $g(k_x, k_y) = k_x k_y$, it follows for

the paraxial condition: $k_x \leq NA_{obj} k$, $k_y \leq NA_{obj} k$. Consequently, the points (k_x, k_y) are enclosed in a circle with the radius $NA_{obj} k$. Now consider an arbitrary point inside the circle. Obviously, the value of $g(k_x, k_y)$ grows if the point moves radially outwards. Thus, the maximum of $g(k_x, k_y)$ is at the circle boundary. Therefore, the relationship $k_x^2 + k_y^2 = NA_{obj}^2 k^2$ describes $g(k_x, k_y)$ as a single variable function $g(k_x) = k_x \sqrt{NA_{obj}^2 k^2 - k_x^2}$. Now the extremum condition $\frac{dg}{dk_x} = 0$ results in $g_{\max} = NA_{obj}^2 k^2 / 2$, which yields the following condition for the discretized scalar $|\nabla_{\perp}^2 \phi|$:

$$|\nabla_{\perp}^2 \phi| \leq \frac{2\sqrt{2} NA_{obj} k}{h}. \quad (16)$$

The inequality (16) is our proposed paraxial constraint on the discretized Laplace operator of the phase. This is only an estimate for the maximum value of the discretized scalar $|\nabla_{\perp}^2 \phi|$ under the assumption of the paraxial wave propagation.

The reason for the dependency of the scalar $|\nabla_{\perp}^2 \phi|_{\max}$ from NA_{obj} is the dependency of the vector $|\nabla_{\perp} \phi|_{\max}$ from NA_{obj} [Inequality (11)] which, as explained before, is the result of the paraxial assumption. The paraxial assumption results in a Fourier transform and consequently the dependency of the transverse spatial frequency from the position in the object plane [Eq. (8)]. However, the dependency of the scalar $|\nabla_{\perp}^2 \phi|_{\max}$ from the spacing h has a different reason; it is the consequence of the discretization of the field (images recorded by the detector). Of course, $|\nabla_{\perp}^2 \phi|_{\max}$ would be independent of the spacing for a continuous field and for the points with continuity property. But, it should be mentioned, that at the points without phase continuity (phase jumping), $|\nabla_{\perp}^2 \phi|$ is not defined or, in other words, infinite. A very interesting point in our view is that in the case of the continuous field, the scalar $|\nabla_{\perp}^2 \phi|$ is infinite at jumping points and thus maybe include no valuable information. The value of $|\nabla_{\perp}^2 \phi|_{\max}$ in the discrete case including the paraxial condition has a finite value given by the inequality (16). In other words, the behavior of the jumping points can be quantified in the discrete case. One mathematical explanation is that in the numerical approaches, the derivative of functions can be approximated by the finite difference method, and automatically the dependency of the spacing h appears. This spacing is the consequence of the finite pixel size of the detector. An example is the step function whose derivative is infinite at the jumping point but, if a discretized step function is considered and the derivative at the jumping point is calculated, it would have a finite value. The smaller the spacing the larger is the derivative value. A similar effect occurs in the evaluation of $|\nabla_{\perp}^2 \phi|_{\max}$. It does not refer only to the physical character of the wave, but also the discretization caused by the detector. The paraxial condition within the discretization considerations provide a value for $|\nabla_{\perp}^2 \phi|_{\max}$, which is an invaluable information used in our algorithm. Moreover, the inequality (16) shows that $|\nabla_{\perp}^2 \phi|_{\max}$ will be infinite if $h \rightarrow 0$. This is an expected result, because of the transition between the discretized and continuous world.

3.2 Implementation of the paraxial constraints

To fulfill the paraxial condition, both constraints for the gradient and Laplace operator of the phase must be considered in the algorithm. These are the necessary constraints to make the set of possible solutions smaller and to prevent non-physical solutions for the phase distribution ϕ . This can be accomplished by establishing the upper limit for $|\nabla_{\perp}\phi|$ and $|\nabla_{\perp}^2\phi|$ as well as fixing their value in the algorithm, if their maximum values are reached already at some points. However, as long as they satisfy both constraints, the values for other points can be changed freely in each iteration.

One further restriction, which has a significant influence on the phase reconstruction, is related to the value of C . Although, from an analytical point of view C can have any arbitrary positive value, the choosing of the numerical value is not trivial. If C is too high, the term $C\nabla_{\perp}^2\phi_n$ is too dominant against the term $k\frac{\partial I}{\partial z}$, which suppresses the crucial information in the intensity derivative. On the other hand, too small values for C cause too high values for $\nabla_{\perp}^2\phi_1$, especially at points with $I_1 = 0$, which may violate again the paraxial condition. Consequently, a reasonable estimation of the constant C is required. Due to the starting phase, $\phi_0(x, y) = 0$, it follows from Eq. (3):

$$(I_1 + C)\nabla_{\perp}^2\phi_1 = -k\frac{\partial I_1}{\partial z}. \quad (17)$$

The condition (16) together with Eq. (17) results in:

$$2\sqrt{2}NA_{obj}|I_1 + C| \geq h\left|\frac{\partial I_1}{\partial z}\right|. \quad (18)$$

The parameter C must satisfy the inequality (16) for arbitrary intensity values including zero:

$$C \geq \frac{h\left|\frac{\partial I_1}{\partial z}\right|_{\max}}{2\sqrt{2}NA_{obj}}. \quad (19)$$

This minimum value will be used in the algorithm.

The values for $|\nabla_{\perp}\phi|$ and $|\nabla_{\perp}^2\phi|$ are fixed, if they reach maximum. However, because $\nabla_{\perp}\phi$ is a vector, the direction of the vector may change. That means in each iteration the value of $|\nabla_{\perp}\phi|$ is checked and if it exceeds $|\nabla_{\perp}\phi|_{\max}$, both components of the vector $\nabla_{\perp}\phi$ will be rescaled so that $|\nabla_{\perp}\phi| = |\nabla_{\perp}\phi|_{\max}$. Thus, the direction of $\nabla_{\perp}\phi$ can change freely but, its magnitude is restricted to its maximum value, according to inequality (11).

4. Results

The approach described in the last sections was applied to the intensity images shown in Figs. 2(a) and 2(b). These two intensity distributions have a distance of $dz = 1\mu m$. To show the strength of the method, we have defined another picture (Fig. 2(c)) as the arbitrary phase distribution of Fig. 2(a). To show the reliability of the method and the fact that it is no longer restricted by the assumption that the intensity must be above zero, we have set 10559 pixels in I_1 to zero (black regions marked with arrows in Fig. 2(a)). The intensity I_2 Fig. 2(b) was calculated from I_1 Fig. 2(a) and the phase distribution Fig. 2(c) by the Angular Spectrum

method [41]. The phase reconstructed from I_1 and I_2 by the proposed algorithm can be seen in Fig. 2(d).

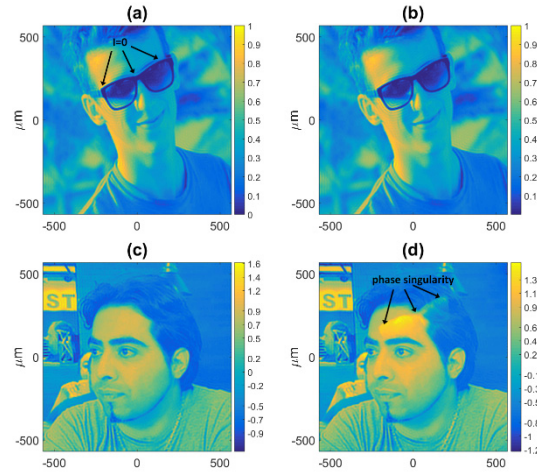


Fig. 2. (a) and (b) are the intensity distributions in the first and the second plane respectively. (c) is the original phase in the first plane and (d) the phase distribution in the first plane reconstructed by the presented method. Both images are quadratic with a width and height of 1.135 mm and the pixel number is 1135 x 1135. The wavelength of the light in the simulation is 633 nm. According to [42], the paraxial condition requires a very small transverse spatial frequency f_t compared to the spatial frequency of the light f ($f = 1/\lambda$). Thus, the maximum transverse spatial frequency of the complex amplitude $u_i = \sqrt{I_i} \exp(i\phi_i)$ was $0.1f$. The used C according to Eq. (19) was 0.057.

In Fig. 2(d) some singularity points can be seen (marked by black arrows). The reason for these phase singularities are the zero values of the corresponding intensity in Fig. 2(a). However, since the phase for an intensity of zero does not make any physical sense, these values are irrelevant. Due to the zero intensity, the corresponding complex amplitude will be zero as well. Thus, the phase value at the singularity points cannot change the complex amplitude (it will always be zero). To evaluate the capability of our method, an inhomogeneous intensity with a non-parallel gradient of intensity and phase and 10559 points with zero intensity have been used. The correlation coefficient and the root-mean-squared error (RMSE) value of the reconstructed phase with the original phase as shown in Figs. 2(c) and 2(d), are 99.2% and 0.069 rad respectively, which were obtained with a total iteration number of 20. Both values were calculated after the reconstructed phase values at the singularity points were put to zero and in the original phase as well. The reason is that at singularity points the reconstructed phase values may be completely different to the original phase. As mentioned before, the difference is physically insignificant, however the difference manipulates the correlation and RMSE value.

Figure 3 shows a similar simulation like Fig. 2 but without any zero-value pixels. As can be expected, the singularity effect disappears. The correlation coefficient and the SNR value are 99.6% and 0.046 rad, respectively.

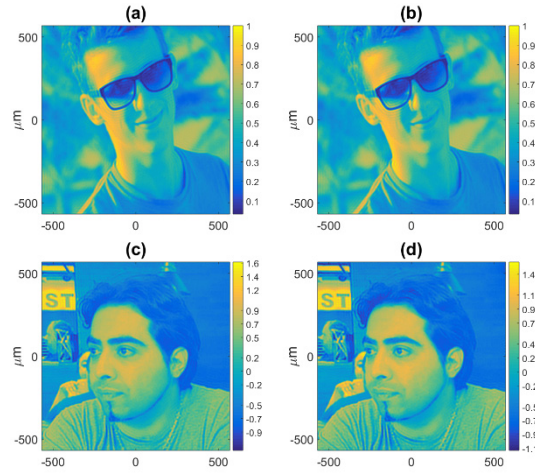


Fig. 3. (a) and (b) are the intensity distributions in the first and the second plane respectively without any zero-value pixels. (c) is the original phase in the first plane and (d) the reconstructed phase distribution in the first plane without any singularity points.

To show the strength of the method in the presence of noise, Fig. 4 presents the correlation coefficients and RMSE values for different signal to noise ratios (SNR) from 10 dB to 40 dB. As can be seen, for the noisy data with an SNR of 15 dB the algorithm provides a reconstructed phase with about 99% correlation coefficient and 0.1 rad for the RMSE. The values improve very fast for increasing SNR. For 40 dB the correlation coefficient and RMSE are 99.3 and 0.06 rad, respectively.

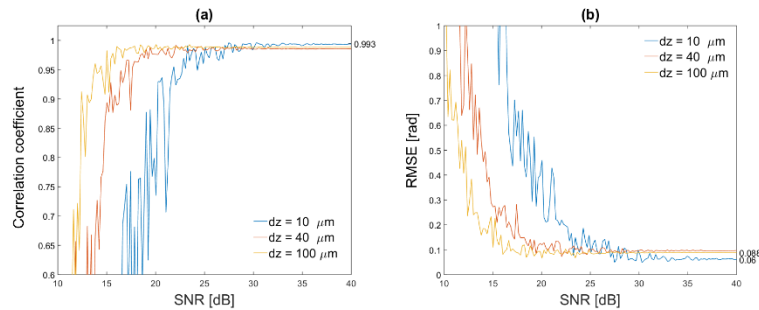


Fig. 4. Correlation coefficient (a) and RMSE values (b) for the image in Fig. 2 in dependence of the SNR.

To suppress the noise influence, for lower SNR the distance dz between the two images has to increase. However, it may cause a worse approximation for $\partial I / \partial z$. The choice of the optimal distance dz has been investigated in many works [38, 43–50]. Especially for the noisy data, it has been demonstrated, that the optimal propagation distance concerning to $\partial I / \partial z$ does not always result in a better phase reconstruction. The phase reconstruction can be improved, if the optimal distance dz related to the phase reconstruction is applied [49]. The aim of this work is the presentation of a new solving approach for the TIE with fewer restrictions thus, here just estimates for the propagation distance dz have been used. We assume that with strategies like optimal dz or multiple measurement approaches for a better $\partial I / \partial z$ estimation, an enhancement of the noise performance can be expected.

One important criteria for the reliability of iterative algorithms is the stability dependence on the iteration number. Especially, for noisy data the convergence of the iterative algorithms

is not trivial. As shown in Figs. 5(a) and 5(b), both correlation coefficient and the RMSE value converge for increasing iteration numbers.

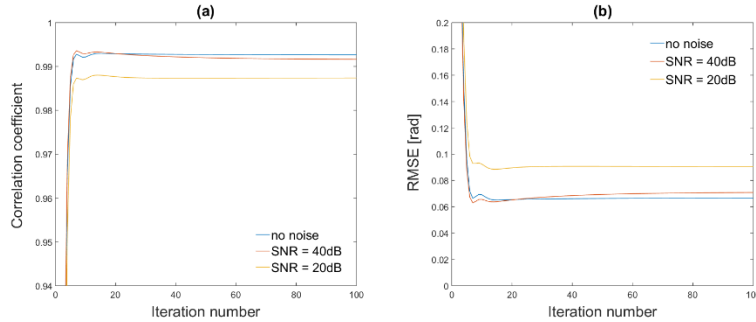


Fig. 5. Stability of the algorithm, (a) and (b) are the correlation and RMSE values respectively for the used image in Fig. 2 dependent on the iteration number.

As can be seen, the noisy data has only influence on the reconstruction quality and not directly on the convergence rate. Furthermore, after only 10 iterations a satisfactory phase reconstruction, of course dependent on the SNR, can be achieved. After around 20 iterations the quality of the reconstruction will almost not change anymore. Consequently, for practical applications and if the implementation time is crucial, the algorithm can provide already satisfactory results for low iteration numbers.

To demonstrate the capability of the algorithm and its implementation efficiency dependent on the number of pixels with zero intensity, different examples have been chosen. Figure 6(a) presents an intensity with zero intensity values distributed in circles with a periodic distribution of the circles.

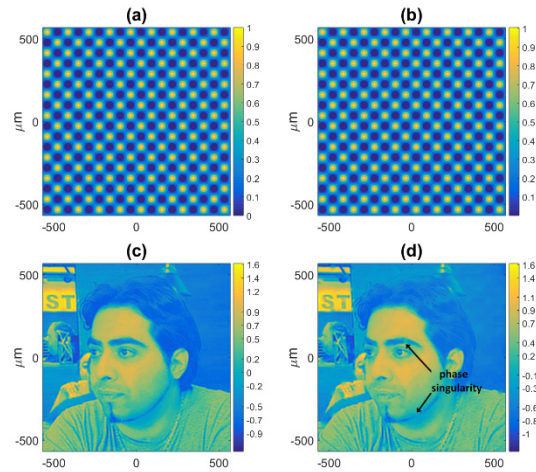


Fig. 6. , circle distribution of the zero intensity values. (a) and (b) Intensity distributions in the first and the second plane, respectively. (c) and (d) Original and reconstructed phase. The amount of zero-value pixels is 15% of the total pixel number.

Figure 6(d) shows the reconstructed phase with the correlation coefficient 99.6% and a RMSE value of 0.046 rad compared to the original phase Fig. 6(c).

Another interesting intensity distribution with zero-value pixels is a grating as shown in Fig. 7.

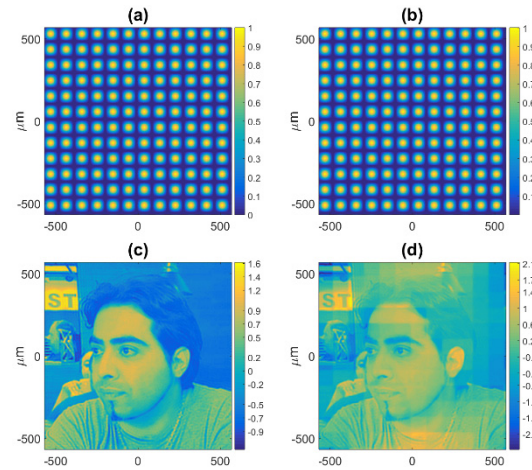


Fig. 7. , grating distribution of the zero intensity values. (a) and (b) Intensity distributions in the first and the second plane, respectively. (c) and (d) Original and reconstructed phase. The amount of zero-value pixels is 3% of the total pixel number. Please note that although the blue area seems to be bigger than in Fig. 6(a) and 6(b), the number of pixels with zero intensity is smaller.

Here the number of pixels with zero intensity is much smaller (3%, compared to 15% in Fig. 6) but, the areas with zero intensity are connected to each other. Thus, the areas with nonzero intensity are surrounded with zero-value pixels. That property leads to a phase shift of the algorithm in the neighboring regions which results in the quadratic structures in Fig. 7(d). This phase shift occurs at the borders of the regions with zero intensity. However, it can be recognized, that in each region the fine structures in the phase distribution are completely reconstructed.

A further example is presented in Fig. 8(a) with zero intensity as a band in the center of the distribution. We use this example as a test for the evaluation of the performance of the algorithm in dependence on the ratio between zero and nonzero pixels.

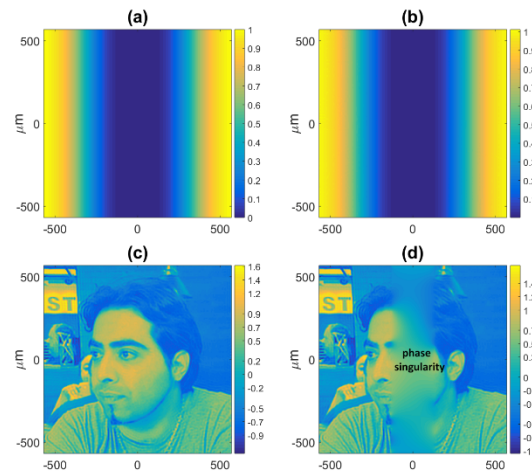


Fig. 8. , band distribution of the zero intensity values. (a) and (b) Intensity distributions in the first and the second plane respectively. (c) and (d) The original and the reconstructed phase. The zero-value pixels are 23% of the total pixel number.

In the two examples presented previously, increasing the size of the areas with zero values reduces the size of the areas with nonzero values. These areas might become so small that, due to diffraction, the paraxial assumption will be violated.

As can be expected, the phase as shown in Fig. 8(d) can be reconstructed at all nonzero areas (compared to Fig. 8(c)). The correlation coefficient and the RMSE value are 0.9996 and 0.03 rad respectively, which show very high reconstruction quality. In the area of the zero-value pixels (the band in the center of Fig. 8(a)), the original phase cannot be retrieved. However, at the singularity points, the phase value is not defined and thus all values are physically equivalent.

Figure 9 presents the effect of the ratio between zero and nonzero pixels on the phase reconstruction for the increasing of the width of the nonzero pixel band in the middle of Fig. 8(a).

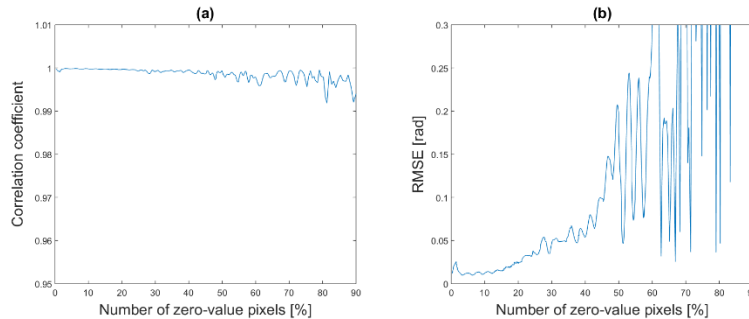


Fig. 9. Dependence of the phase retrieval quality from the number of zero-value pixels, (a) and (b) are the correlation and RMSE values, respectively for the intensity distribution in Fig. 8(a) with an increased zero intensity area.

Both, correlation coefficient and the RMSE value are very high for values up to 40%. However, they decrease for higher pixel numbers.

5. Conclusion

An iterative method for removing the usual intensity restrictions in solving the TIE has been proposed. Contrary to the conventional Fourier-based methods, the intensity must not be homogeneous, and the intensity and phase gradients must not be parallel. Moreover, the method presented here allows zero values for the intensity. The validity and capability of the method, has been verified by investigations of the noise influence. The algorithm provides very high quality results for SNRs higher than 25dB. The influence of the number of the zero-value pixels has been investigated as well. The reliability of the algorithm for zero-nonzero pixel ratios up to 40% has been shown. Moreover, the stability of the algorithm has been verified. For iteration numbers of about 10, we get a very high quality reconstruction. As demonstrated, the presented method can be used as a feasible and high reliable approach for solving the TIE without the usual solving assumptions in similar Fourier-based methods.

Funding

Niedersächsisches Vorab (NL – 4 Project “QUANOMET”) and German Research Foundation (DFG SCHN 716/13-1).

Acknowledgments

We like to thank Janosch Meier, Ali Dorostkar and Melvin Küster for very fruitful discussions and Okan Özdemir for his support, Stefan Preußler for the proofreading of the manuscript and his support and invaluable suggestions during the writing of the paper and Naghmeh Akbari for the creation of Fig. 1.